1 Jonas literature study 04/22

1. [?]
   (a) They devise a minimax optimal test that relies on the so-called k-sparse largest eigenvalue of the empirical covariance matrix. It captures the largest amount of empirical variance explained by any k of the original variables. The test tries to detect sparse principal components.
   (b) Classical PCA produces inconsistent estimators of the directions that explain the most variance. [?] (and more in paper)
   (c) Johnstone Johnstone2001 introduces the spiked covariance model which provides a natural setting for statistical problems. The model relies on the assumption that there exists a small number of directions that explain most of the variance. Data is drawn from a multi-Gaussian distribution with mean zero and covariance matrix given by $I + \theta vv^T$, v is a unit norm sparse vector.
2. [?]
   (a) Too complex don't discuss
3. [?]
   (a) Some initial reduction in dimensionality is desirable before applying PCA.
   (b) The initial reduction in dimensionality is best achieved by working in a basis in which the signals have a sparse representation.
   (c) Overall too dense to read properly
4. [?]
   (a) Sparse PCA is employed to identify a small number foe interpretable directions that represent the data succinctly.
   (b) Regular PCA works well when $n >> p$, but starts faltering when $p > n$ (see also [?]).
   (c) Since the meaningfulness of the components depends on the characteristics of the problem (sample size, dimensionality, sparsity level and signal-to-noise ratio).
   (d) [?] shows that his estimator for the SPCA attains the minimax rate of convergence over a certain Gaussian class of distributions, provided that the sparsity parameter k is treated as a fixed constant. [?, ?] allows for k to vary depending on the sample size n of the problem. These papers settle the question of sparse principal component estimation. However they are not computable in polynomial time.

(e) [**?**] made improvement in this computability problem by testing versus an alternative hypothesis.

(f) They themselves also try to improve computability.

5. [**?**]

   (a) Study SPCA in high dimensions where $p >> n$ and introduce to complementary notions: row sparsity and column sparsity.

   (b) Problem: whether SPCA methods can optimally estimate the subspace spanned by the leading eigenvectors, that is, the principal subspace of variation.

   (c) [**?**] proposed a sparse principal subspace estimator based on iterative thresholding, and derived its rate of convergence under a spiked covariance model. similar to [**?**].

   (d) The use of the spiked covariance model is warranted by two reasons. First of all, it simplifies analyses and enables the exploitation of special properties of the multivariate Gaussian distribution. The second is that it excludes the possibility of the variables having equal variances.

6. [**?**]

   (a) Present a penalized matrix decomposition (PMD) framework for computing a rank-K approximation of a matrix. The matrix X is approximated as $\hat{X} = \sum_{k=1}^{K} d_k u_k v_k^T$ where d_k, u_k and v_k minimize the squared Frobenius norm of $X - \hat{X}$ subject to penalities on u_k and v_k. They find that if they apply ℓ_1-penalities on v_k it yields an efficient algorithm for the SCoTLASS proposal [**?**].